CONFERENCE 27 – 30 November 2017
EXHIBITION 28 – 30 November 2017
BITEC, Bangkok, Thailand
SA2017.SIGGRAPH.ORG

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).
SA '17 Courses, November 27-30, 2017, Bangkok, Thailand ACM 978-1-4503-5403-5/17/11. 10.1145/3134472.3134507

Tutorial on Information Theory in Visualization

presenters: Mateu Sbert, Han-Wei Shen, Ivan Viola
co-authors: Min Chen, Anton Bardera, Miquel Feixas

Introduction to Information Theory

Mateu Sbert

University of Girona (Spain)

Tianjin University (China)

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

OVERVIEW

- Introduction
- Information measures
 - entropy, conditional entropy
 - mutual information
- Information channel
- Relative entropy
- Mutual information decomposition
- Inequalities
- Information bottleneck method
- Entropy rate

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

INTRODUCTION (1)

- Claude Elwood **Shannon**, 1916-2001
- "**A mathematical theory of communication**", Bell System Technical Journal, July and October, 1948
- The significance of Shannon's work
- Transmission, storage and processing of information
- Applications: physics, computer science, mathematics, statistics, biology, linguistics, neurology, computer vision, etc.

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

INTRODUCTION (2)

- Certain quantities, like **entropy** and **mutual information**, arise as the answers to fundamental questions in communication theory
- **Shannon entropy** is the ultimate data compression or the expected length of an optimal code
- **Mutual information** is the communication rate in presence of noise
- Book: T.M. Cover and J.A. Thomas, Elements of Information Theory, Wiley, 1991, 2006

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

INTRODUCTION (3)

- **Shannon** introduced two fundamental concepts about "**information**" from the communication point of view
 - information is **uncertainty**
 - information source is modeled as a random variable or a random process
 - probability is employed to develop the information theory
 - information to be transmitted is **digital**
 - Shannon's work contains the first published use of "bit"
- Book: R.W. Yeung, Information Theory and Network, Springer, 2008

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (1)

- Random variable X taking values in an alphabet X

$$X : \{x_1, x_2, \dots, x_n\}, p(x) = \Pr\{X = x\}, p(X) = \{p(x), x \in X\}$$

- Shannon entropy $H(X)$, $H(p)$: uncertainty, information, homogeneity, uniformity

$$H(X) = - \sum_{x \in X} p(x) \log p(x) \equiv - \sum_{i=1}^n p(x_i) \log p(x_i)$$

- information associated with x : $-\log p(x)$; base of logarithm: 2; convention: $\log 0 = 0$; unit: bit: uncertainty of the toss of an ordinary coin

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (2)

- Examples:

- Entropy of a fair coin toss:

$$H(X) = -(1/2) \log(1/2) - (1/2) \log(1/2) = 1 \text{ bit}$$

- Entropy of a fair die toss:

$$H(X) = -(1/6) \log(1/6) - (1/6) \log(1/6) \dots = 2.58 \text{ bits}$$

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

INFORMATION MEASURES (3)

- **Properties** of Shannon entropy

- $0 \leq H(X) \leq \log|X|$
- binary entropy: $H(X) = -p \log p - (1-p) \log(1-p)$
- $H(p_1, p_2, p_3) = H(p_1, p_2 + p_3) + (p_2 + p_3) H\left(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}\right)$

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (4)

$$H(0.001, 0.002, 0.003, 0.980, 0.008, 0.003, 0.002, 0.001) = 0.190$$

$$H(0.010, 0.020, 0.030, 0.800, 0.080, 0.030, 0.020, 0.010) = 1.211$$

$$H(0.200, 0.050, 0.010, 0.080, 0.400, 0.010, 0.050, 0.200) = 2.314$$

$$H(0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125) = 3.000$$

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (5)

- Discrete random variable Y in an alphabet Y

$$Y : \{y_1, y_2, \dots, y_n\}, p(y) = \Pr\{Y = y\}$$

- Joint entropy

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y)$$

- Conditional entropy

$$\begin{aligned} H(Y | X) &= \sum_{x \in X} p(x) H(Y | x) = - \sum_{x \in X} p(x) \sum_{y \in Y} p(y | x) \log p(y | x) \\ &= - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(y | x) \end{aligned}$$

SA2017.SIGGRAPH.ORG

INFORMATION CHANNEL

- Communication or information channel $X \rightarrow Y$

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (6)

- Mutual information $I(X;Y)$: shared information, correlation, dependence, information transfer

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y | X) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x \in X} p(x) \sum_{y \in Y} p(y | x) \log \frac{p(y | x)}{p(y)} \end{aligned}$$

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES (7)

- Relationship between information measures

$$0 \leq H(X|Y) \leq H(X)$$

$$H(X,Y)$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$H(X)$$

$$H(Y)$$

$$I(X;Y) \leq H(X)$$

$$I(X;Y)$$

$$I(X;Y) = I(Y;X) \geq 0$$

$$H(X,Y) = H(X) + H(Y) - I(X;Y)$$

Yeung's book: Chapter 3 establishes a one-to-one correspondence between Shannon's information measures and set theory. A number of examples are given to show how the use of information diagrams can simplify the proofs of many results in information theory.

SA2017.SIGGRAPH.ORG

INFORMATION MEASURES EXAMPLE

As an example, we consider the joint distribution $p(X, Y)$ represented in Fig. 1.3.*left*. The marginal probability distributions of X and Y are given by $p(X) = \{0.25, 0.25, 0.5\}$ and $p(Y) = \{0.375, 0.625\}$, respectively. Thus, $H(X) = -0.25 \log 0.25 - 0.25 \log 0.25 - 0.5 \log 0.5 = 1.5$ bits, $H(Y) = -0.375 \log 0.375 - 0.625 \log 0.625 = 0.954$ bits, and $H(X, Y) = -0.125 \log 0.125 - 0.125 \log 0.125 - 0.25 \log 0.25 - 0 \log 0 - 0 \log 0 - 0.5 \log 0.5 = 1.75$ bits.

$p(X, Y)$	$\mathcal{Y}$		$p(X)$
	y_1	y_2	
x_1	0.125	0.125	0.25
x_2	0.25	0	0.25
x_3	0	0.5	0.5
$p(Y)$	0.375	0.625	
$H(X, Y) = 1.75$			

$p(Y X)$	$\mathcal{Y}$		$H(Y x \in \mathcal{X})$
	y_1	y_2	
x_1	0.5	0.5	$H(Y x_1) = 1$
x_2	1	0	$H(Y x_2) = 0$
x_3	0	1	$H(Y x_3) = 0$
$H(Y X) = 0.25$			

$$\begin{aligned}
 H(Y|X) &= \sum_{i=1}^3 p(x_i) H(Y|X = x_i) \\
 &= 0.25 H(Y|X = x_1) + 0.25 H(Y|X = x_2) + 0.5 H(Y|X = x_3) \\
 &= 0.25 \times 1 + 0.25 \times 0 + 0.5 \times 0 = 0.25 \text{ bits.}
 \end{aligned}$$

INFORMATION MEASURES (8)

- Normalized mutual information: different forms

$$\frac{I(X;Y)}{H(X,Y)}$$

$$\frac{I(X;Y)}{H(X) + H(Y)}$$

$$\frac{I(X;Y)}{\min\{H(X), H(Y)\}}$$

$$\frac{I(X;Y)}{\max\{H(X), H(Y)\}}$$

- Information distance

$$H(X|Y) + H(Y|X)$$

SA2017.SIGGRAPH.ORG

RELATIVE ENTROPY

- Relative entropy, informational divergence, Kullback-Leibler distance
 $D_{KL}(p,q)$: how much p is different from q (on a common alphabet X)

$$D_{KL}(p,q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$$

- convention: $0 \log 0/q = 0$ and $p \log p/0 = \infty$
- $D_{KL}(p,q) \geq 0$
- it is not a true metric or "distance" (non-symmetric, triangular inequality is not fulfilled)
- $I(X;Y) = D_{KL}(p(X,Y), p(X)p(Y))$

SA2017.SIGGRAPH.ORG

MUTUAL INFORMATION

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y | X) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x \in X} p(x) \sum_{y \in Y} p(y | x) \log \frac{p(y | x)}{p(y)} \end{aligned}$$

$$D_{KL}(p,q) = \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$$

$$I(X;Y) = D_{KL}(p(X,Y), p(X)p(Y))$$

SA2017.SIGGRAPH.ORG

MUTUAL INFORMATION DECOMPOSITION

- Information associated with x

$$I(X;Y) = \sum_{x \in X} p(x) \left[\sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)} \right] = \sum_{x \in X} p(x) (H(Y) - H(Y|x))$$

$$I_1(x;Y) = \sum_{y \in Y} p(y|x) \log \frac{p(y|x)}{p(y)}$$

$$I_2(x;Y) = H(Y) - H(Y|x)$$

[DeWeese]

$$I_3(x;Y) = \sum_{y \in Y} p(y|x) I_2(X;y)$$

[Butts]

$$I(X;Y) = \sum_{x \in X} p(x) I_k(x;Y)$$

$k = 1, 2, 3$

INEQUALITIES

- **Data processing inequality:** if $X \rightarrow Y \rightarrow Z$ is a Markov chain, then

$$I(X;Y) \geq I(X;Z)$$

No processing of Y can increase the information that Y contains about X , i.e., further processing of Y can only increase our uncertainty about X on average

- **Jensen's inequality:** a function $f(x)$ is said to be convex over an interval (a,b) if for every x_1, x_2 in (a,b) and $0 \leq \lambda \leq 1$

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

SA2017.SIGGRAPH.ORG

JENSEN-SHANNON DIVERGENCE

- From the concavity of entropy, **Jensen-Shannon divergence**

$$JS(\pi_1, \dots, \pi_N; p_1, \dots, p_N) = H\left(\sum_{i=1}^N \pi_i p_i\right) - \sum_{i=1}^N \pi_i H(p_i) \geq 0$$

[Burbea]

- $JS(\pi_1, \dots, \pi_N; p_1, \dots, p_N) = \sum_{i=1}^N \pi_i D_{KL}\left(p_i, \sum_{i=1}^N \pi_i p_i\right)$
- $JS(p(x_1), \dots, p(x_n); p(Y | x_1), \dots, p(Y | x_n)) = I(X; Y)$

SA2017.SIGGRAPH.ORG

INFORMATION CHANNEL, MI AND JS

- Communication or information channel $X \rightarrow Y$

$$JS(p(x_1), \dots, p(x_n); p(Y | x_1), \dots, p(Y | x_n)) = I(X; Y)$$

INFORMATION BOTTLENECK METHOD (1)

- Tishby, Pereira and Bialek, 1999
- To look for a compressed representation of X which maintains the (mutual) information about the relevant variable Y as high as possible

SA2017.SIGGRAPH.ORG

INFORMATION BOTTLENECK METHOD (2)

- **Agglomerative information bottleneck method**: clustering/merging is guided by the minimization of the loss of mutual information

$$I(X;Y) \geq I(\hat{X};Y)$$

- Loss of mutual information

$$I(X;Y) - I(\hat{X};Y) = p(\hat{x})JS(p(x_1)/p(\hat{x}), \dots, p(x_m)/p(\hat{x}); p(Y|x_1), \dots, p(Y|x_m))$$

$$\text{where } p(\hat{x}) = \sum_{k=1}^m p(x_k)$$

[Slonim]

- The quality of each cluster is measured by the Jensen-Shannon divergence between the individual distributions in the cluster

INFORMATION CHANNEL AND IB

- Communication or information channel $X \rightarrow Y$

$$I(X;Y) - I(\hat{X};Y) =$$

$$p(\hat{x})JS(p(x_1)/p(\hat{x}), p(x_2)/p(\hat{x}); p(Y|x_1), p(Y|x_2))$$

$$p(\hat{x}) = p(x_1) + p(x_2)$$

SA2017.SIGGRAPH.ORG

EXAMPLE: ENTROPY OF AN IMAGE

- The information content of an image is expressed by the Shannon entropy of the (normalized) intensity histogram

- The entropy disregards the spatial contribution of pixels

EXAMPLE: IMAGE PARTITIONING (1)

- Information channel $X \rightarrow Y$ defined between the **intensity histogram** and the **image regions**

$$X \xrightarrow{p(Y|X)} Y$$

b_i = number of pixels of bin i ;
 r_j = number of pixels of region j
 N = total number of pixels

SA2017.SIGGRAPH.ORG

EXAMPLE: IMAGE PARTITIONING (2)

information bottleneck method

information gain

$$I(X;Y) - I(\hat{X};Y) = p(\hat{x})JS(p(x_1)/p(\hat{x}), p(x_2)/p(\hat{x}); p(Y|x_1), p(Y|x_2))$$

at each step, increase of $I(X;Y)$ = decrease of $H(X|Y)$

$$H(X) = I(X;Y) + H(X|Y)$$

EXAMPLE: IMAGE PARTITIONING (3)

$$MIR = \frac{I(\hat{X};Y)}{I(X;Y)} ; \text{ number of regions ; \% of regions}$$

0.1; 13; 0.00

0.2; 64; 0.02

0.3; 330; 0.13

0.4; 1553; 0.59

0.0; 5597; 2.14

1; 234238; 89.35

0.9; 129136; 49.26

0.8; 67291; 25.67

0.7; 34011; 12.97

0.6; 15316; 5.84

SA2017.SIGGRAPH.ORG

ENTROPY RATE

- Shannon entropy

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log p(x)$$

- Joint entropy

$$H(X^L) = - \sum_{x^L \in \mathcal{X}^L} p(x^L) \log p(x^L)$$

- Entropy rate or information density

$$h = \lim_{L \rightarrow \infty} \frac{H(X^L)}{L}$$

$$= \lim_{L \rightarrow \infty} (H(X^L) - H(X^{L-1}))$$

SA2017.SIGGRAPH.ORG

Viewpoint metrics and applications

Mateu Sbert

University of Girona, Tianjin University

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

VIEWPOINT SELECTION

- Automatic selection of the most informative viewpoints is a very useful focusing mechanism in visualization
- It can guide the viewer to the most interesting information of the data set
- A selection of most informative viewpoints can be used for a virtual walkthrough or a compact representation of the information the data contains
- Best view selection algorithms have been applied to computer graphics domains, such as scene understanding and virtual exploration, N best views selection , image-based modeling and rendering, mesh simplification, molecular visualization, and camera placement
- Information theory measures have been used as viewpoint metrics since the work of Vazquez et al. [2001], see also [Sbert et al. 2009]

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

THE VISUALIZATION PIPELINE

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

DIRECT VOLUME RENDERING (DVR)

- Volume dataset is considered as a transparent gel with light travelling through it

- classification maps primitives to graphical attributes
- shading (illumination) models shadows, light scattering, absorption...
 - usually absorption + emission optical model
- compositing integrates samples with optical properties along viewing rays

Transfer function
definition

Local or global
illumination

Both realistic and illustrative rendering

VIEWPOINT SELECTION

- Takahashi et al. 2005
 - Evaluation of viewpoint quality based on the visibility of extracted isosurfaces or interval volumes.
 - Use as viewpoint metrics the average of viewpoint entropies for the extracted isosurfaces.

$$E_i^{iso}(v) = \frac{-1}{\log(m_i + 1)} \sum_{j=0}^{m_i} \frac{A_{ij}}{S} \log \frac{A_{ij}}{S}$$

$$E^{iso}(v) = \sum_{i=1}^n \frac{E_i^{iso}(v)}{n}$$

SA2017.SIGGRAPH.ORG

VIEWPOINT SELECTION

- Takahashi et al. 2005

Best and worst views of interval volumes extracted from a data set containing simulated electron density distribution in a hydrogen atom

SA2017.SIGGRAPH.ORG

VIEWPOINT SELECTION

- Bordoloi and Shen 2005
 - Best view selection: use entropy of the projected visibilities distribution

$$H(v) = - \sum_{i=1}^n q_i(v) \log q_i(v)$$

- Representative views: cluster views according to Jensen-Shannon similarity measure

$$JS(\frac{1}{2}, \frac{1}{2}; q(v_1), q(v_2)) = H(\frac{1}{2}q(v_1) + \frac{1}{2}q(v_2)) - H(q(v_1)) - H(q(v_2))$$

VIEWPOINT SELECTION

- Bordoloi and Shen 2005

Best (two left) and worst (two right) views of tooth data set

Four representative views

VIEWPOINT SELECTION

- Ji and Shen 2006
 - Quality of viewpoint v , $u(v)$, is a combination of three values

$$u(v) = \beta_1 \text{opacity}(v) + \beta_2 \text{color}(v) + \beta_3 \text{curvature}(v)$$

VIEWPOINT SELECTION

- Mühler et al. 2007
 - Semantics-driven view selection. Entropy, between other factors, used to select best views.
 - Guided navigation through features assists studying the correspondence between focus objects.

VISIBILITY CHANNEL

- Viola et al. 2006, Ruiz et al. 2010

$$p(v) = \frac{vis(v)}{\sum_{i \in \mathcal{V}} vis(i)}$$

$$p(z|v) = \frac{vis(z|v)}{vis(v)}$$

$p(V)$	$p(Z V)$				
$p(v_1)$	$p(z_1 v_1)$	$p(z_2 v_1)$	$\dots$	$p(z_m v_1)$	
$p(v_2)$	$p(z_1 v_2)$	$p(z_2 v_2)$	$\dots$	$p(z_m v_2)$	
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	
$p(v_n)$	$p(z_1 v_n)$	$p(z_2 v_n)$	$\dots$	$p(z_m v_n)$	
$p(Z)$	$p(z_1)$	$p(z_2)$	$\dots$	$p(z_m)$	

$$p(z) = \sum_{v \in \mathcal{V}} p(v)p(z|v)$$

- How a viewpoint sees the voxels
- Mutual information

$$I(V; Z) = \sum_{v \in \mathcal{V}} p(v) \sum_{z \in \mathcal{Z}} p(z|v) \log \frac{p(z|v)}{p(z)} = \sum_{v \in \mathcal{V}} p(v) I(v; Z)$$

- Viewpoint mutual information (VMI): $I(v; Z) = \sum_{z \in \mathcal{Z}} p(z|v) \log \frac{p(z|v)}{p(z)}$

SA2017.SIGGRAPH.ORG

REVERSED VISIBILITY CHANNEL

- Ruiz et al. 2010

$$p(v|z) = \frac{p(v)p(z|v)}{p(z)}$$

$p(Z)$	$p(V Z)$			
$p(z_1)$	$p(v_1 z_1)$	$p(v_2 z_1)$	$\cdots$	$p(v_m z_1)$
$p(z_2)$	$p(v_1 z_2)$	$p(v_2 z_2)$	$\cdots$	$p(v_m z_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$p(z_n)$	$p(v_1 z_n)$	$p(v_2 z_n)$	$\cdots$	$p(v_m z_n)$
$p(V)$	$p(v_1)$	$p(v_2)$	$\cdots$	$p(v_m)$

- How a voxel “sees” the viewpoints
- Mutual information

$$I(Z; V) = \sum_{z \in Z} p(z) \sum_{v \in V} p(v|z) \log \frac{p(v|z)}{p(v)} = \sum_{z \in Z} p(z) I(z; V)$$

- Voxel mutual information (VOMI): $I(z; V) = \sum_{v \in V} p(v|z) \log \frac{p(v|z)}{p(v)}$

SA2017.SIGGRAPH.ORG

VOMI MAP COMPUTATION

Transfer
function

VISIBILITY CHANNEL

- Viola et al. 2006
- Adding importance to VMI for viewpoint navigation with focus of interest. Objects instead of voxels

$$I(v; O) = D_{KL}(p(O|v) || p(O))$$

$$I'(v; O) = D_{KL}(p(O|v) || p'(O)) = \sum_{o \in O} p(o|v) \log \frac{p(o|v)}{p'(o)}$$

VOMI APPLICATIONS

- Interpret VOMI as ambient occlusion
 - $AO(z) = 1 - \overline{I(z; V)}$
 - Simulate global illumination
 - Realistic and illustrative rendering
 - Color ambient occlusion
 - $CAO_{\alpha}(z; V) = \sum_{v \in \mathcal{V}} \left(p(v|z) \log \frac{p(v|z)}{p(v)} \right) (1 - C_{\alpha}(v))$

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

VOMI APPLICATIONS

- Interpret VOMI as importance
 - Modulate opacity to obtain focus+context effects emphasizing important parts
- “Project” VOMI to viewpoints to obtain informativeness of each viewpoint
 - $INF(v) = \sum_{z \in Z} p(v|z)I(z; V)$
 - Viewpoint selection

SA2017.SIGGRAPH.ORG

VOMI AS AMBIENT OCCLUSION MAP

Original

Ambient Occlusion, Landis 2002
Vicinity shading, Stewart 2003

Obscurances, Iones et al. 98

VOMI

VOMI APPLIED AS AMBIENT OCCLUSION

- Ambient lighting term

- Additive term to local lighting

Original

Vicinity shading,
Stewart 2003

VOMI

COLOR AMBIENT OCCLUSION

CAO map

CAO map
with contours

CAO maps with contours
and color quantization

SA2017.SIGGRAPH.ORG

OPACITY MODULATION

Original

Modulated to emphasize skeleton

Original

Modulated to
emphasize ribs

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

VIEWPOINT SELECTION

- VMI versus Informativeness

Min VMI

Max VMI

Min VMI

Max VMI

Max INF

Min INF

Max INF

Min INF

REFERENCES

- T.M. Cover and J.A. Thomas. Elements of Information Theory. Wiley, 1991, 2006
- R.W. Yeung. Information Theory and Network. Springer, 2008
- M.R. DeWeese and M. Meister. How to measure the information gained from one symbo., Network: Computation in Neural Systems, 10, 4, 325-340, 1999
- D.A. Butts. How much information is associated with a particular stimulus?. Network: Computation in Neural Systems, 14, 177-187, 2003
- J. Burbea and C.R. Ra. On the convexity of some divergence measures based on entropy functions. IEEE Transactions on Information Theory, 28, 3, 489-495, 1982
- Noam Slonim and Naftali Tishby. Agglomerative Information Bottleneck. NIPS, 617-623, 1999

SA2017.SIGGRAPH.ORG

REFERENCES

- Imre Csiszár and Paul C. Shields. Information Theory and Statistics: A Tutorial. Communications and Information Theory, 1, 4, 2004
- Pere P. Vazquez, Miquel Feixas, Mateu Sbert, and Wolfgang Heidrich. Viewpoint selection using viewpoint entropy. In Proceedings of Vision, Modeling, and Visualization 2001 , pages 273-280, Stuttgart, Germany, November 2001.
- M. Sbert, M. Feixas, J. Rigau, M. Chover, I. Viola. Information Theory Tools for Computer Graphics. Morgan and Claypool Publishers, 2009
- Bordoloi, U.D. and Shen, H.-W. (2005). View selection for volume rendering. In IEEE Visualization 2005 , pages 487-494
- Ji, G. and Shen, H.-W. (2006). Dynamic view selection for time-varying volumes. Transactions on Visualization and Computer Graphics , 12(5):1109-1116

SA2017.SIGGRAPH.ORG

REFERENCES

- Mühler, K., Neugebauer, M., Tietjen, C. and Preim, B. (2007). Viewpoint selection for intervention planning. In Proceedings of Eurographics/ IEEE-VGTC Symposium on Visualization, 267-274
- Ruiz, M., Boada, I., Feixas, M., Sbert, M. (2010). Viewpoint information channel for illustrative volume rendering. Computers & Graphics , 34(4):351-360
- Takahashi, S., Fujishiro, I., Takeshima, Y., Nishita, T. (2005). A feature driven approach to locating optimal viewpoints for volume visualization. In IEEE Visualization 2005 , 495-502
- Viola, I, Feixas, M., Sbert, M. and Gröller, M.E. (2006). Importance-driven focus of attention. IEEE Transactions on Visualization and Computer Graphics , 12(5):933-940

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

CONFERENCE 27 – 30 November 2017
EXHIBITION 28 – 30 November 2017
BITEC, Bangkok, Thailand
SA2017.SIGGRAPH.ORG

THANKS FOR YOUR ATTENTION!

Mateu Sbert

University of Girona, Tianjin University

Section 2

Visualization & Information Theory

author: Min Chen, University of Oxford

presenter: Mateu Sbert, University of Girona, Tianjin University

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC**, Bangkok, Thailand

ENTROPY

- Random variable (**alphabet**)
 - X
- It takes values (**letters**)
 - $x_1, x_2, \dots, x_m$
- Probability mass function
 - $p(x_i)$
- Entropy (**uncertainty**)
 - $$\mathbf{H}(X) = -\sum_i^m p(x_i) \log_2 p(x_i)$$

SA2017.SIGGRAPH.ORG

THE ROLE OF A THEORETIC FRAMEWORK

Facts

Wisdom

Theory

AN INFORMATION-THEORETIC VIEW OF VISUALIZATION

M. Chen and H. Jänicke, *An information-theoretic framework for visualization*,
IEEE Transactions on Visualisation and Computer Graphics, 2010

A General Visualization System

A General Communication System

SOURCE ENCODING AND CHANNEL ENCODING

CHAPTER 1 ■ Basic Concepts of Information Theory	1
1.1 ENTROPY	2
1.2 RELATIVE ENTROPY AND MUTUAL INFORMATION	5
1.3 DECOMPOSITION OF MUTUAL INFORMATION	6
1.4 ENTROPY RATE	8
1.5 INEQUALITIES	9
1.5.1 Jensen's Inequality	9
1.5.2 Jensen-Shannon Inequality	10
1.5.3 Data Processing Inequality	11
1.6 INFORMATION BOTTLENECK METHOD	
1.7 CONTINUOUS CHANNEL	
CHAPTER 2 ■ Visualization and Information Theory	15
2.1 INFORMATION-THEORETIC MEASURES IN VISUALIZATION	18
2.1.1 Alphabets and Letters	18
2.1.2 Quantifying Visual Information	19
2.1.3 Information Sources and Communication Channels	23
2.1.3.1 Information Sources	24
2.1.3.2 Channels	24
2.1.4 Visual Coding in Noiseless Channels	26
2.1.5 Visual Coding in Noisy Channels	28
2.2 INFORMATION-THEORETIC LAWS FOR VISUALIZATION	29
2.3 INFORMATION-THEORETIC PROCESS OPTIMIZATION	33
2.3.1 Data Processing Inequality	34

2.3.2 Transformation of Alphabets in a Visualization Process	35
2.3.3 Cost-benefit Measures for Visualization Processes	37
2.3.4 Examples of Cost-benefit Analysis	42
2.3.4.1 Interaction in Visualization	43
2.3.4.2 Disseminative Visualization	46
2.3.4.3 Observational Visualization	47
2.3.4.4 Analytical Visualization	48
2.3.4.5 Model-developmental Visualization	48
2.4 INFORMATION-THEORETIC LINKS: VISUALIZATION AND PERCEPTION	49
2.4.1 Visual Multiplexing	49
2.4.1.1 Multiplexing in Communication	49
2.4.1.2 Multiplexing in Visualization	50
2.4.1.3 Multiplexing and Information Theoretic Measures	55
2.4.2 Information-theoretic Quality Metrics	57
2.4.2.1 Differentiability of Glyphs	58
2.4.2.2 Hamming Distance	60
2.4.2.3 Quasi-Hamming Distance for Glyph Design	61
2.4.3 User Studies in the Information-theoretic Framework	63
2.5 SUMMARY	64

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

SHANNON'S DEFINITION OF INFORMATION

SA2017.SIGGRAPH.ORG

SHANNON'S DEFINITION OF INFORMATION

SA2017.SIGGRAPH.ORG

AN ALPHABET AND ITS LETTERS

- English alphabet
- All English prefixes
- All English words
- All sentences in a text corpus
- ...
- All published Siggraph papers
- ...

A, B, C, ..., X, Y, Z
bio, geo, pre, pro, ..., un
a, ..., silicosis, ..., titin, ...
...

45 letters

pneumonoultramicroscopicsilicovolcanoconiosis

189,819 letters

a word or a formula?

AN EXAMPLE OF VISUALIZATION (OR VA) PROCESS

MUTUAL INFORMATION (SHARED UNCERTAINTY)

alphabet
 $\mathbb{X}$

alphabet
 $\mathbb{Y}$

$$I(X; Y) = \sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} p(x, y) \log_2 \frac{p(x, y)}{p_X(x)p_Y(y)}$$

DATA PROCESSING INEQUALITY

Decreasing of Mutual Information

$$I_1 \geq I_2 \geq \dots \geq I_{L-1} \geq I_L$$

DATA PROCESSING INEQUALITY

“No clever manipulation of data can improve the inferences that can be made from the data”

[Cover and Thomas, 2006]

DATA PROCESSING INEQUALITY WITH “BIG DATA”

Can one trust a decision with a very small amount of mutual information with the data?

DPI IS NOT UBIQUITOUS

- Markov chain conditions
 - Closed coupling: (X, Y), (Y, Z)
 - X and Z are conditionally independent
- What if one of the conditions is broken?
- In visual analytics, both conditions are usually broken.

$$I(X; Y) \overset{\times}{\geq} I(X; Z)$$

SA2017.SIGGRAPH.ORG

DATA PROCESSING WITH SOFT KNOWLEDGE

$x \in \mathbb{X}$ is a piece of soft knowledge about the background and context of the data and processing.

DATA PROCESSING WITH HUMANS-IN-THE-LOOP

DATA PROCESSING WITH HUMANS-IN-THE-LOOP

*All possible decisions
under different conditions*

a) totally data-driven

b) totally instinct-driven

c) data-informed

*d) due to unknown or
uncontrollable factors*

TRANSFORMATION

COST-BENEFIT RATIO

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

COST-BENEFIT RATIO

Benefit

$$\frac{\mathcal{B}(F_s)}{\mathcal{C}(F_s)} = \frac{\mathcal{H}(Z_s) - \mathcal{H}(Z_{s+1}) - \mathcal{D}_{KL}(Z'_s || Z_s)}{\mathcal{C}(F_s)}$$

Cost

Cost

SA2017.SIGGRAPH.ORG

KULLBACK-LEIBLER DIVERGENCE

$$D_{KL}(Z'|Z) = \sum_{(z=z') \in Z} p(z') \log_2 \frac{p(z')}{q(z)}$$

Solomon Kullback
1907-1994

Richard Leibler
1914-2003

SA2017.SIGGRAPH.ORG

COMPOSITION

SA2017.SIGGRAPH.ORG

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

THREE SPACES AND THREE MEASURES

M. Chen and H. Jänicke, **An information-theoretic framework for visualization**, *IEEE Transactions on Visualisation and Computer Graphics*, 2010

- Entropy of Input Data Space: $H(X)$
- Visualization Capacity: $V(G)$
- Display Capacity: D

$$\text{Visual Mapping Ratio (VMR)} = \frac{V(G)}{H(X)}$$

$$\text{Information Loss Ratio (ILR)} = \frac{\max(H(X) - V(G), 0)}{H(X)}$$

$$\text{Display Space Utilization (DSU)} = \frac{V(G)}{D}$$

SA2017.SIGGRAPH.ORG

EXAMPLE OF $V(G)$

- Entropy of the Data Alphabet: $H(Z) = -\sum_{t=0}^{64} \sum_{i=0}^{255} \frac{1}{256} \log_2 \frac{1}{256} = 512$
- $V(G) = H(Z)$
- Binary Pixel Plot: D
 - 4x4 pixels per bit $\rightarrow D: 2^9 \times 2^4 = 2^{13}$ bits
- Time Series Plot: D
 - Minimal 256x64 pixels $\rightarrow D: 2^8 \times 2^6 = 2^{14}$ bits
- The more compact, the better?
- Cost? Reconstructability?

SA2017.SIGGRAPH.ORG

$H(X) = 512$ bits
 $V(G) = 512$ bits
 $D = 16382$ bits
 $VMR = 1$
 $ILR = 0$
 $DSU = 0.03$

$H(X) = 512$ bits
 $V(G) = 384$ bits
 $D = 4096$ bits
 $VMR = 0.75$
 $ILR = 0.25$
 $DSU = 0.09$

(a) evenly distributed p

INFORMATION LOSS RATIO (ILR)

- Display Space Restriction
 - 64x64 pixels
- Evenly distributed probability mass function
- Linear visual mapping
- ILR is a probabilistic measure about
 - a data space X , not an instance x_i

SA2017.SIGGRAPH.ORG

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
- 3. Measurement, Explanation, and Prediction**
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC**, Bangkok, Thailand

information loss:
25%

M. Chen and H. Jänicke, **An information-theoretic framework for visualization**, *IEEE Transactions on Visualisation and Computer Graphics*, 2010

INFORMATION LOSS RATIO (ILR)

- Display Space Restriction
 - 64x64 pixels
- Evenly distributed probability mass function
- Linear visual mapping
- ILR is a probabilistic measure about
 - a data space X , not an instance x_i

SA2017.SIGGRAPH.ORG

NON-UNIFORM DISTRIBUTION

information loss:
25%

(a) evenly distributed p

information loss:
25.8%

(b) unevenly distributed p

- Linear visual mapping

NON-UNIFORM DISTRIBUTION

- Non-linear visual mapping

information loss:
25%

(a) evenly distributed p

information loss:
25.8%

(b) unevenly distributed p

information loss:
22.6%

(c) 4 regional mappings

information loss:
0%

(d) logarithmic plot

NON-UNIFORM DISTRIBUTION

- Logarithmic visual mapping

D: $p = \frac{1}{8}$

C: $p = \frac{1}{8}$

B: $p = \frac{1}{4}$

A: $p = \frac{1}{2}$

$p = \frac{1}{2^k}$

$p = \frac{1}{2^k}$

$p = \frac{1}{2^{k-1}}$

.....

$p = \frac{1}{2^3}$

$p = \frac{1}{2^2}$

$p = \frac{1}{2}$

A PREDICTION?

- Display capacity D is limited.
- Data space: noticeable non-uniform distribution: $H(X) \ll H_{\max}(X)$
- Visualization capacity $V(G)$: a visual representation exhibiting a similar non-uniform distribution of space requirement can reduce information loss.
- Can adjacency matrices be improved for trees?

<http://hci.stanford.edu/jheer/files/zoo/>

<http://en.wikipedia.org/wiki/Treemapping>

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
- 4. Example: Visual Multiplexing**
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC**, Bangkok, Thailand

MAP OVERLAY AND ITS GENERALIZATION IN VISUALIZATION

(a) Robertson *et al.* [RFF*08]
Type B

(b) Everts *et al.* [EBR109]
Type C

(c) Bair, House [BH07]
Types C, D, G

(d) Treavett, Chen [TC00]
Type D, J

(e) Collins *et al.* [CPC09]
Types C, E

(f) Guo *et al.* [GXY12]
Type E

(g) Kindlmann, Westin [KW06]
Type F

(h) Ware [War09]
Types G, J

(i) Chen *et al.* [CPL*11]
Types C, G, J

(j) Saito *et al.* [SMY*05]
Type C, H

(k) Drocourt *et al.* [DBS*11]
Type H

(l) Ware, Plumlee [WP13]
Type I

(m) Viola *et al.* [VFSG06]
Type E, J

(n) Correa *et al.* [CSC06]
Type H, J

(o) Botchen *et al.* [BBS*08]
Types E, G, H, J

(p) Maguire *et al.* [MRSS*12]
Type J

VISUAL MULTIPLEXING

M. Chen, S. Walton, K. Berger, J. Thiyyagalingam, B. Duffy,
H. Fang, C. Holloway, and A. E. Trefethen,
Visual multiplexing, *Computer Graphics Forum*, 2014.

- Frequency-division multiplexing (FDM)?
- Time-division multiplexing (TDM)?
- Space-division multiplexing (SDM)?
- Code-division multiplexing (CDM)?

10 TYPES OF VISUAL MULTIPLEXING

(a) Type A: Partition a space

(b) Type B: Partition a time period

(c) Type C: Introduce partial occlusion

(d) Type D: Use a 'hollow' visual channel

(e) Type E: Introduce translucent occlusion

(f) Type F: Use an integrated visual channel

(g) Type G: Depict a continuous field

(h) Type H: Shift a visual channel

(i) Type I: Use periodic motion

(j) Type J: Assume a priori knowledge

(k) Type J (continued): Acquired knowledge

(l) Type J (continued): Visual language

TYPE C: INTRODUCE PARTIAL OCCLUSION

TYPE G: DEPICT A CONTINUOUS FIELD

TYPE H: SHIFT A VISUAL CHANNEL

HOW ABOUT WHEN EVERY PIXEL IS USED?

$$\frac{V(G)}{D} \ll 1$$

SYSTEMATIC APPLICATION

- C: Introduce Partial Occlusion
- E: Introduce Translucent Occlusion
- G: Depict a Continuous Field
- J: Assume A Priori Knowledge

R. P. Botchen, S. Bachthaler, F. Schick, M. Chen, G. Mori, D. Weiskopf and T. Ertl, **Action-based multi-field video visualization**, *IEEE Transactions on Visualization and Computer Graphics*, 2008.

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
- 5. Example: Error Detection and Correction**
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
7. Summary

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

FOUR LEVELS OF VISUALIZATION

1. Disseminative Level

$O(1)$

This is “a”!

2. Observational Level

$O(n)$

“a”, “b”, “c”, ... what, when, where?

3. Analytical Level

$O(n^k)$

Are “a”, “b”, “c” related? Why?

4. Model-developmental Level

$O(k^n)$, $O(n!)$

How does “a” lead to “b”?

SA2017.SIGGRAPH.ORG

A COMPOSITE WORKFLOW

OBSERVATIONAL VISUALIZATION

P. A. Legg, D. H. S. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths and M. Chen,
MatchPad: Interactive Glyph-Based Visualization for Real-Time Sports Performance Analysis, *Computer Graphics Forum*, 2012

Input Video

0

OBSERVATIONAL

BBC News - Wales' Rugby World Cup team using Swansea Uni app - Mozilla Firefox

File Edit View History Bookmarks Tools Help

BBC News - Wales' Rugby World Cup te...

http://www.bbc.co.uk/news/uk-wales-15105398

BBC Mobile News Sport Weather iPlayer TV Radio More Search BBC News

NEWS WALES

Home World UK England N.Ireland Scotland Wales Business Politics Health Education Sci/Environment Technology Entertainment & Arts

Wales Politics North West North East Mid South West South East Newyddion

29 September 2011 Last updated at 13:28

Wales' Rugby World Cup team using Swansea Uni app

Welsh coaches are using technology developed for them by experts at Swansea University to improve their match analysis at the Rugby World Cup.

Coach Warren Gatland and his backroom team are bombarded with statistics about scrums, line-outs and tackles during each game.

Now they are using an 'app' to simplify the information they need and understand the team's performance.

It also allows the coaches to review video of key moments during play.

The Welsh team employs three analysts whose job it is to collate data about all aspects of each game from the set pieces and restarts to tackles made or missed.

Dr Philip Legg of Swansea University's college of engineering and department of computer science said the problem was coaches were suffering from an "information overload."

The university was approached by the Welsh Rugby Union (WRU) and has developed an app it calls the MatchPad which runs on an Apple iPad.

It produces a visual timeline during the game so analysts and coaches can review video and additional detail on the events they are most interested in simply by pressing an icon.

The portability of the device means they can access the information in the analysis box, in the changing rooms, or even at pitch-side.

Dr Legg said: "During each game the team analysts are busy recording each of the events that happen.

"They look at each scrum, line out, restart, possession won and lost and tackles.

"They collect so much information - that's the basic problem and the app just tries to simplify.

The MatchPad was tested by the Welsh team at Cardiff's Millennium Stadium

Related Stories

Wales 'strongest' Euro cup hope

Fiji 0-66 Wales

Wales 81-7 Namibia

Top Stories

Italy's Berlusconi faces key vote

Lloyds hit by cost of PPI claims

Boxing heavyweight Frazier dies

Sarkozy called Israeli PM 'liar' NEW

Doctor guilty over Israeli death

Features & Analysis

Staging da Vinci

How the National Gallery's blockbuster came together

Patriotic purchase

Why aren't people told to buy British?

Break down

Who is really profiting from the high cost of petrol and oil?

Brain teaser

Debunking the myths about our minds

Most Popular

Shared Read Video/Audio

Sarkozy called Israeli PM 'liar' 1

Boxing heavyweight Frazier dies 2

Giant asteroid to pass near Earth 3

Prostitutes found in Mexico jail 4

Why aren't people being told to buy British? 5

OUTLINE

1. Data Intelligence — a big picture
2. Visualization — a small picture
3. Measurement, Explanation, and Prediction
4. Example: Visual Multiplexing
5. Example: Error Detection and Correction
6. Example: Process Optimization
- 7. Summary**

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

Information Theory Taxonomy	Relevance in Visualization
Fundamental Concepts	A possible mathematical framework that underpins the subject of visualization.
Major Quantities and Properties	Quantitative measurements about the data and visualization space, and the relationship between input and output of a process or subsystem at different stages of a visualization pipeline.
Entropy	Measuring information content (see Section 5.1); salience in visualization.
Mutual Information	Uncertainty reduction in visualization (see Section 5.3); information-assisted visualization.
Major Theorems	Many theorems can be used to explain visualization phenomena and events.
Information balance (conservation law)	Given two visualizations, A and B, the amount of information about A contained in B is the same as that about B in A; overview + detail visualization; multi-view visualization.
Data processing inequality	After visual mapping, the visualization cannot contain more information than the original data (see Section 6.2); Information cannot be recovered after being degraded by some processes or subsystems in a visualization pipeline.
Channel Types	Providing a theoretical basis for classifying visualization subsystems (see Section 6).
Noiseless channel	Not common in practical visualization pipelines (see Section 3).
Noisy channel	Most visualization processes and subsystems can be affected by noise (see Section 3).
Channel Capacity	It can be adapted to define the maximum amount of information that can be visualized or displayed (see Sections 5.1 and 6)
Redundancy	Efficiency of a visual mapping; Error detection and correction (see Section 8).
Source Coding (for Noiseless Channels)	Inspiration for developing new data abstraction and visual encoding techniques.
Coding Schemes	Applicable, for example, to the following visualization algorithms:
Entropy coding (e.g., Huffman, arithmetic coding)	Logarithmic plots (see Section 7.1); importance-based visualization; information-assisted visualization; magic lens; illustrative deformation.

CONFERENCE 27 – 30 November 2017
EXHIBITION 28 – 30 November 2017
BITEC, Bangkok, Thailand
SA2017.SIGGRAPH.ORG

Volume Visualization

Information Theory in Visualization

Ivan Viola

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 **EXHIBITION** 28 – 30 November 2017 **BITEC, Bangkok, Thailand**

AGENDA

- Time-Varying Data
- Level-of-Detail Selection
- IsoSurfaces
- Splitting
- Transfer function specification

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

VIEW SELECTION FOR VOLUME DATA

- 3D scalar fields over time
- Viewpoint quality: visibility of voxels
- Measure: Joint viewpoint entropy

$$\begin{aligned}
 H(X) &= H(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \\
 &= H(X_{t_1}) + H(X_{t_2} | X_{t_1}) + \dots + H(X_{t_n} | X_{t_n-1})
 \end{aligned}$$

Bordoloi and Shen 2005]

IMPORTANCE-DRIVEN VISUALIZATION

- Quantify data-block importance using entropy within a block and mutual information wrt. neighboring blocks

$$H(X) - I(X, Y) = H(X|Y)$$

- Concatenating data-block importance
- over time
- Clustering time-importance curves
- Selecting spatial region corresponding to the curve set

SA2017.SIGGRAPH.ORG

IMPORTANCE-DRIVEN VISUALIZATION

IMPORTANCE-DRIVEN VISUALIZATION

[Wang et al. 2008]

SA2017.SIGGRAPH.ORG

EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

IMPORTANCE-DRIVEN VISUALIZATION

APH.ORG

[Wang et al. 2008]

MULTIRESOLUTION VOLUMES

- Distortion (D) and Contribution (C) characteristics of a multiresolution block
- Level-of-Detail quality evaluated via Entropy measure

$$p_i = \frac{C_i \cdot D_i}{\sum_{i=1}^M C_i \cdot D_i}$$

$$D_i = \sum_{j=0}^7 d_{ij} + \max(D_j |_{j=0}^7)$$

$$C_i = \mu \cdot l \cdot a \cdot v$$

- Join and split optimizations driven by entropy changes

$$d_{ij} = \sigma_{ij} \frac{\mu_i^2 + \mu_j^2 + c_1}{2\mu_i\mu_j + c_1} \frac{\sigma_i^2 + \sigma_j^2 + c_2}{2\sigma_i\sigma_j + c_2}$$

[Wang and Shen 2006]

MULTIRESOLUTION VOLUMES

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

ISOSURFACE SIMILARITY MAPS

- Quantifying the most significant iso-surfaces in volume
- Compare iso-surfaces through evaluating mutual information of their distance volume
 - X and Y are independent: $I(X, Y) = 0$
 - X and Y are identical: $I(X, Y) = H(X) = H(Y)$

SA2017.SIGGRAPH.ORG

[Bruckner and Möller 2010]

ISOSURFACE SIMILARITY MAPS

- Normalized measure

$$\hat{I}(X, Y) = \frac{2I(X, Y)}{H(X) + H(Y)}$$

SA2017.SIGGRAPH.ORG

ISOSURFACE SIMILARITY MAPS

$$SD(i) = \frac{1}{|V|} \sum_{j=1}^{|V|} SM(i, j)$$

ISOSURFACE SIMILARITY MAPS

- Selection of characteristic iso-surfaces

MULTIMODAL SIMILARITY MATRIX

[Haidacher et al. 2011]

SIMILARITY-BASED EXPLODED VIEWS

- A two step process is proposed to automatically obtain the partitioning planes:
 - Explosion axis: selection of the most structured view
 - Partitioning of the data: slices are grouped according to the maximization of a similarity criterion

SIMILARITY-BASED EXPLODED VIEWS

- Structured View measured through Entropy Rate
measure of the randomness or unpredictability of a system

SIMILARITY-BASED EXPLODED VIEWS

$$H(X^L) = - \sum_{x^L \in X^L} p(x^L) \log p(x^L)$$

- Explosion Axis:
Joint entropy of L vector

- Entropy rate represents the average information content per symbol in a stochastic process

$$h = \lim_{L \rightarrow \infty} \frac{H(X^L)}{L} = \lim_{L \rightarrow \infty} (H(X^L) - H(X^{L-1}))$$

- For a set of viewpoints we perform ray casting to compute entropy rate for each ray
- Summing up entropy rates per direction quantifies how much structural change particular direction exhibit

SIMILARITY-BASED EXPLODED VIEWS

- Bottom-up grouping: group the most similar slices or slabs through normalized mutual information
degree of similarity or shared information between two slices or slabs

SIMILARITY-BASED EXPLODED VIEWS

TRANSFER FUNCTIONS FOR SCALAR FIELDS

SA2017.SIGGRAPH.ORG

[Ruiz et al. 2011]

TRANSFER FUNCTIONS FOR SCALAR FIELDS

- Measuring visibility $vis(b|v)$ needed data-values b from viewpoint v
- Constructing a communication channel between viewpoints and bins with transition probability matrix and marginal probabilities

$$p(b|v) = vis(b|v) / vis(v)$$

$$vis(v) = \sum_{b \in \mathcal{B}} vis(b|v)$$

- Average projected $p(v) = vis(v) / \sum_{i \in \mathcal{V}} vis(i)$

$$p(b) = \sum_{v \in \mathcal{V}} p(v) p(b|v)$$

$p(X)$	$p(Y X)$				
$p(x_1)$	$p(y_1 x_1)$	$p(y_2 x_1)$	...	$p(y_m x_1)$	.
$p(x_2)$	$p(y_1 x_2)$	$p(y_2 x_2)$	...	$p(y_m x_2)$	
⋮	⋮	⋮	⋮	⋮	
⋮	⋮	⋮	⋮	⋮	
$p(x_n)$	$p(y_1 x_n)$	$p(y_2 x_n)$	...	$p(y_m x_n)$	
	$p(y_1)$	$p(y_2)$	...	$p(y_m)$	$p(Y)$

PH.ORG

30 November 2017 BITEC, Bangkok, Thailand

[Ruiz et al. 2011]

TRANSFER FUNCTIONS FOR SCALAR FIELDS

- Target distribution $q(B)$:
 - Uniform for all data values
 - Proportional to bin size
 - Proportional to value
 - Proportional to gradient magnitude
 - Proportional to importance
- Informational divergence (KL-distance): quantifies the distance between true $p(B)$ and target $q(B)$ distributions

$$D_{KL}(p(B), q(B)) = \sum_{b \in \mathcal{B}} p(b) \log \frac{p(b)}{q(b)}$$

TRANSFER FUNCTIONS FOR SCALAR FIELDS

- Iterative Transfer Function Design
- Minimize informational divergence between the average projected visibility distribution from all viewpoints and a target distribution
- Transforming the opacity vector
- Optimizer: Steepest Gradient Descent

$$A = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$$

$$A_{t+1} = A_t - \nabla D_{KL}(p||q)$$

SA2017.SIGGRAPH.ORG

TRANSFER FUNCTIONS FOR SCALAR FIELDS

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

TRANSFER FUNCTIONS FOR SCALAR FIELDS

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

REFERENCES

- [Bordoloi and Shen 2005] U. Bordoloi and H.W. Shen: View Selection for Volume Rendering, In Proceedings of IEEE Visualization, 2005
- [Ji and Shen 2006] G. Ji and H.W. Shen: Dynamic View Selection for Time-Varying Volumes, In IEEE TVCG (Proceedings of IEEE Visualization), 2006
- [Wang et al. 2008] C. Wang, H. Yu, K.-L. Ma: Importance-Driven Time-Varying Data Visualization, IEEE TVCG (Proceedings of IEEE Visualization), 2008
- [Wang and Shen 2006] C. Wang and H.W. Shen: LOD Map - A Visual Interface for Navigating Multiresolution Volume Visualization, In IEEE TVCG (Proceedings of IEEE Visualization), 2006
- [Bruckner and Möller 2010] S. Bruckner and T. Möller: Isosurface Similarity Maps, In EG CGF (Proceedings of EuroVis), 2010

SA2017.SIGGRAPH.ORG

CONFERENCE 27 – 30 November 2017 EXHIBITION 28 – 30 November 2017 BITEC, Bangkok, Thailand

REFERENCES

- [Ruiz et al. 2008] M. Ruiz, I. Viola, I. Boada, S. Bruckner, M. Feixas, M. Sbert: Similarity-based Exploded Views, In *Springer LNCS (Proceedings of Smart Graphics)*, 2008
- [Ruiz et al. 2011] M. Ruiz, A. Bardera, I. Boada, I. Viola, M. Feixas, M. Sbert: Automatic Transfer Functions based on Informational Divergence, *IEEE TVCG (Proceedings of IEEE Visualization)*, 2010
- [Haidacher et al. 2008] M. Haidacher, S. Bruckner, A. Kanitsar, and M.E. Gröller: Information-based Transfer Functions for Multimodal Visualization, In *Proceedings of EG VCBM*, 2008
- [Bramon et al. 2012] R. Bramon, I. Boada, A. Bardera, J. Rodriguez, M. Feixas, J. Puig, M. Sbert: Multimodal Data Fusion based on Mutual Information, *IEEE TVCG*, 2012
- [Haidacher et al. 2011] M. Haidacher, S. Bruckner, M. E. Gröller: Volume Analysis Using Multimodal Surface Similarity, *IEEE TVCG (Proceedings of IEEE Visualization)*, 2011

SA2017.SIGGRAPH.ORG

THANK YOU

viola@cg.tuwien.ac.at

CONFERENCE 27 – 30 November 2017
EXHIBITION 28 – 30 November 2017
BITEC, Bangkok, Thailand
SA2017.SIGGRAPH.ORG

Flow Visualization

Tutorial on Information Theory in Visualization

Han-Wei Shen
The Ohio State University

Entropy for Scientific Data

- A data set can be considered as a random variable
- Each data point can be considered as an outcome of the random variable
- We can estimate the information content of the whole data set or of local regions

SA2017.SIGGRAPH.ORG

Distributions from Scientific Data

Scalar Distributions

Vector Distributions

Feature Distributions

State Transitions

SA2017.SIGGRAPH.ORG

Data Sets with Multiple Variables

- Assuming your data set contains two variables X and Y
- You want to know the relationship between X and Y
- You can calculate the conditional entropy, mutual information, etc between these two variables
- Some of the metrics can be used as the 'information distance' between two variables

SA2017.SIGGRAPH.ORG

Entropy for Multiple Variables

- Joint Entropy

$$H(X, Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y)$$

- Conditional Entropy

$$H(X|Y) = \sum_{y \in \mathcal{Y}} p(y) H(X|Y = y) = - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(x, y) \log p(x|y)$$

- Mutual Information

$$I(X; Y) = H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

Relations of Entropy Measures

SA2017.SIGGRAPH.ORG

Evaluating Visualization

Vector Field Analysis

- Concept

- Treat the vector field as a data source that generates vector orientation as outcome
- The more diversified are the vector orientations, the more information is contained in the vector field

- Measurement

- Estimate the distribution of the vector orientation
- Compute the entropy of this distribution as the measurement

Vector field

Polar Histogram

SA2017.SIGGRAPH.ORG

Information in Vector Fields

SA2017.SIGGRAPH.ORG

Entropy Field and Seeding

Measure the entropy around each point's neighborhood

Vector Field

Entropy field: higher value means more information in the corresponding region

Entropy-based seeding: Places streamlines on the region with high entropy

SA2017.SIGGRAPH.ORG

Evaluation of Visualization

SA2017.SIGGRAPH.ORG

Information Comparison Between Data/Visualization

Vector Field **X**

Streamlines **Y**

Conditional entropy $H(X|Y)$:

The information in **X** not represented by **Y**

An effective visualization should represent most information in the data,
i.e. $H(X|Y)$ should be small

SA2017.SIGGRAPH.ORG

Conditional Entropy Field and Seeding

Measure the under-represented information in local regions

Streamlines

Conditional entropy field

Conditional-entropy-based seeding: Place more seeds on regions with higher under-represented information

SA2017.SIGGRAPH.ORG

Result

1st iteration: Entropy-based seeding

2nd iteration: Cond.-entropy-based seeding

Conditional entropy

When conditional entropy converges

SA2017.SIGGRAPH.ORG

View-dependent Flow Visualization

- Goal: create a clear view of important features in 3D flow fields
- Issue: occlusion of the flow features
- Approaches
 - Evaluate flow field in screen space with information theory
 - Place streamline to highlight salient flow features with reduced occlusion

SA2017.SIGGRAPH.ORG

Maximal Entropy Projection (MEP)

MEP: Project the entropy field to the screen via Maximal Intensity Projection (MIP)

- Measure the maximal entropy visible to each pixel
- Store the sampled entropy and depth in the MEP Framebuffer

SA2017.SIGGRAPH.ORG

MEP-based Streamline Placement

- Highlight salient flow features
- Reduce occlusion to these features

Streamline Evaluation

Input Streamlines

MEP Framebuffer

SA2017.SIGGRAPH.ORG

Streamline Evaluation

Input Streamlines

Streamlines w/ less occlusion
to the MEP Framebuffer

MEP Framebuffer

SA2017.SIGGRAPH.ORG

Streamline Evaluation

Input Streamlines

Streamlines w/ less occlusion
to the MEP Framebuffer

MEP Framebuffer

Streamlines that occluded
the MEP Framebuffer

SA2017.SIGGRAPH.ORG

MEP-based Streamline Placement

SA2017.SIGGRAPH.ORG

Streamline Statistical Feature Descriptors

- Each streamline is represented as one or more distributions of feature measures such as curvature, curl and torsion

SA2017.SIGGRAPH.ORG

Streamline Statistical Feature Descriptors

- Problem of 1D histograms
 - The order of features is not preserved in the final histogram

A streamline with only one high curvature zone

A streamline with two high curvature zone

SA2017.SIGGRAPH.ORG

Streamline Statistical Feature Descriptors

- Solution: 2D Histograms
 - Decompose the streamline into a fixed number of segments
 - Create 1D histogram of appropriate quantity for each segment
 - Stack the 1D histograms to form a 2D histogram which preserve the order between segments

Streamline Decomposition

- An iterative segmentation algorithm
- Recursively divide into segments until:
 - The difference in the 1D histograms between two halves is smaller than a threshold
 - Streamline segment is too short to be further segmented

SA2017.SIGGRAPH.ORG

Measure Similarity Between Two Streamlines

- Compute similarity between the 2D histograms of two streamlines
 - As two streamline have different number of segments,
 - Apply **Dynamic Time Warping (DTW)** to find an optimal mapping between segments
 - For each pair of segments,
 - Use **Earth Mover's Distance** to measure the distance of their 1D histograms

Similarity-based Streamline Query (Hurricane Isabel Data Set)

- Streamlines having similar features as the one selected by the user are displayed to highlight features in the data
- Histograms based on Curvature and Torsion are used to answer query in this particular case

Hurricane Isabel

User selected target

User selected target

Top 400 matches
SA2017.SIGGRAPH.ORG

Top 200 matches

Similarity-based Streamline Query (Solar Plume Data Set)

- Query response using curvature and torsion based histograms

Solar Plume

User selected streamline

Top 200 matches

User selected streamline

Top 20 matches

SA2017.SIGGRAPH.ORG

Similarity-based Streamline Query

(Ocean Data Set)

SA2017.SIGGRAPH